blood

Analysis of somatic hypermutation in X-HIgM syndrome reveals specific deficiencies in mutational targeting

Nancy S. Longo, Patricia L. Lugar, Sule Yavuz, Wen Zhang, Peter H. L. Krijger, Daniel E. Russ, Dereje D. Jima, Sandeep S. Dave, Amrie C. Grammer and Peter E. Lipsky

Information about reproducing this article in parts or in its entirety may be found online at:
http://bloodjournal.hematologylibrary.org/misc/rights.dtl#repub_requests

Information about ordering reprints may be found online at:
http://bloodjournal.hematologylibrary.org/misc/rights.dtl#reprints

Information about subscriptions and ASH membership may be found online at:
http://bloodjournal.hematologylibrary.org/subscriptions/index.dtl

IMMUNOBIOLOGY

Analysis of Somatic Hypermutation in X-HIgM Syndrome Reveals Specific Deficiencies in Mutational Targeting

Nancy S. Longo[1], Patricia L. Lugar[1]*, Sule Yavuz[1]*, Wen Zhang[1], Peter H. L. Krijger[1], Daniel E. Russ[2], Dereje D. Jima[3], Sandeep S. Dave[34], Amrie C. Grammer[1] and Peter E. Lipsky[1]

[1]Autoimmunity Branch, National Institute of Arthritis, and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland 20892-1560

[2]Division of Computational Bioscience, Center for Information Technology, National Institutes of Health, Bethesda, Maryland 20892-5624

[3]Duke Institute for Genome Sciences and Policy, Duke University, Box 3382, Durham, North Carolina 27710, USA

[4]Department of Medicine, Duke University Medical Center, Durham, North Carolina 27710, USA

Corresponding author:

Dr. Peter E. Lipsky

Autoimmunity Branch

National Institute of Arthritis and Musculoskeletal and Skin Diseases

9000 Rockville Pike

Building 10, Room 6D47C

Bethesda, MD 20892-1560

Tel: (240) 751-2225

Fax: (301) 402-2209

Email: peterlipsky@comcast.net

*The current address for Dr. Patrica Lugar is the Department of Medicine, Duke University, Durham, NC, 27708 and for Dr. Yavuz is the University of Marmara, Dept. of Rheumatology, Istanbul, Turkey

Abbreviations used in this paper: AICDA, activation induced cytidine deaminase; BER, base excision repair; CDR, complementarity determining region; FR, framework region; GC, germinal center; MMR, mismatch repair; MZ, marginal zone; ND, normal donor; NER, nucleotide excision repair; NHEJ, nonhomologous end joining; PB, peripheral blood; PCNA, proliferating cell nuclear antigen; POL, polymerase; R, replacement; RPA, replication protein A; SHM, somatic hypermutation; S, silent; TLS, translesion synthesis; TNF, tumor necrosis factor; Ts, transition; Tv, transversion; UNG, uracil-DNA glycosylase.

Abstract

Subjects with X-linked hyper IgM Syndrome (X-HIgM) have a markedly reduced frequency of CD27$^+$ memory B cells and their Ig genes have a low level of somatic hypermutation (SHM). To analyze the nature of SHM in X-HIgM, we sequenced 209 nonproductive and 926 productive Ig heavy chain genes. In nonproductive rearrangements that were not subjected to selection, as well as productive rearrangements, most of the mutations were within targeted RGYW, WRCY, WA or TW motifs (R=purine, Y=pyrimidine and W= A or T). However, there was significantly decreased targeting of the hypermutable G in RGYW motifs. Moreover, the ratio of transitions:transversions was markedly increased compared to normal. Microarray analysis documented that specific genes involved in SHM, including activation-induced cytidine deaminase (AICDA) and uracil-DNA glycosylase (UNG2), were upregulated in normal germinal center (GC) B cells, but not induced by CD40 ligation. Similar results were obtained from light chain rearrangements. These results indicate that in the absence of CD40-CD154 interactions, there is a marked reduction in SHM and specifically, mutations of AICDA-targeted G residues in RGYW motifs along with a decrease in transversions normally related to UNG2 activity.

Introduction

Encounter with antigen and CD40-CD154 co-stimulation activates naïve B cells to undergo a germinal center (GC) reaction[1,2] during which somatic hypermutation (SHM) introduces point mutations in the Ig variable segment involved in antigen recognition and, thus is an important mechanism for increasing avidity maturation and repertoire diversity.[3] In X-HIgM syndrome, B cells are intrinsically normal but fail to receive the required CD40-CD154 co-stimulation to undergo GC reactions and are unable to generate T cell-dependent memory responses.[4,5] However, a modestly mutated IgD$^+$CD19$^+$ population has been reported in X-HIgM patients with null CD154 mutations.[5,6] The explanation for the appearance of B cells with mutated Ig genes in the absence of GCs remains controversial. It has been suggested that the IgD$^+$CD19$^+$ mutated cells that reside within the CD27$^+$ subset may be increased in frequency and represent a circulating splenic marginal zone population,[7] although others have reported no elevation of circulating CD27$^+$ cells in X-HIgM.[8] Alternatively, antigen activated B cells may start mutating before a GC reaction.[9,10] Thus, the mutations found in X-HIgM B cells could represent the pre-GC or an extra-GC component of SHM.

Whereas the mutational pattern of GC experienced normal peripheral memory B cells has been well characterized,[11] the spectrum of mutations that develop in the absence of GCs is not known. The goal of this analysis, therefore, was to compare the features of normal peripheral blood (PB) B cell mutations with those that occur in the absence of CD40-CD154 interactions and GCs to gain further insight into the nature of SHM and the B cells that undergo SHM in X-HIgM.

Activation-induced cytidine deaminase (AICDA) plays an essential role in SHM.[12,13] Although AICDA was originally considered GC B cell-specific, AICDA expression can be

induced by signaling through various receptors that activate NF-κB, such as the CD40, TLR and LPS pathways.[12,14,15] It can be upregulated in human PB B cells stimulated *in vitro* with combinations of anti-IgM or anti-CD40 with IL-21 stimulation[16] and is expressed by extrafollicular large proliferating B cells.[17,18] AICDA is expressed in Peyer's patch lymphocytes from CD40[-/-] mice[19] as well as intestinal lymphocytes from CD40 deficient humans[20] and is most likely expressed in activated X-HIgM B cells that are intrinsically normal. Notably, AICDA is not constitutively expressed by human B cells residing in the marginal zone of the human spleen.[21] This raises the question as to the nature and patterns of mutations in X-HIgM CD27[+] B cells, especially if they possibly originate from marginal zone cells in response to T-independent antigens.[7,22]

The hallmark of SHM is the increased percentage of mutations in hypermutable RGYW (underlined residue is the most frequently mutated) motifs.[23] Subsequently, the mutation frequency of WRCY motifs, the reverse complement of RGYW, was noted to be similar to that of RGYW[24] and more recently this motif was recognized as the target of AICDA.[13] The G:C mutation spectrum in SHM is thought to be the direct consequence of AICDA deamination of C residues in WRC motifs[13] on the non-transcribed strand[25] during transcription[26] as well as on the transcribed strand.[27-29]

Despite the central role of AICDA in G/C mutations, during SHM nearly half of the mutations target A and T residues.[11] The A:T mutation spectrum is associated with mismatch repair utilizing MSH2-MSH6[28,30-33] and exonuclease 1[34] which normally provides high fidelity repair that presumably becomes mutagenic in GC B cells with the recruitment of error-prone polymerase eta (POL η)[35,36] targeting WA motifs. Pol η deficiency results in a mutational

PAGE 2

pattern lacking targeting to WA motifs.[37,38] Therefore, Pol η appears to account for most WA mutations that can comprise as much as 25% of total nucleotide substitutions.[37]

Finally, the action of uracil–DNA glycosylase (UNG) appears to account for the introduction of transversions. This enzyme deletes AICDA-induced uracil and initiates translesion synthesis (TLS) involving mechanisms that generate transversions that don't occur when high fidelity polymerase recognizes the dU:dG mismatch as dT.[39]

The present study involved a detailed mutation analysis of PB B cells from 4 young and 3 adult X-HIgM patients to characterize the mutation patterns that develop in the absence of CD154 stimulation and GC reactions. We were particularly interested in mutations in nonproductive heavy and light chain rearrangements because these are not influenced by selection.[11] Because AICDA and POL η generate mutations in specific nucleotide motifs and UNG2 is involved in the generation of transversions, we focused on analyzing these specific features of mutational targeting to determine the frequency and nature of the mutations occurring in X-HIgM B cells. The results indicate that mutational targeting in X-HIgM are significantly different than that noted in normal memory B cells populations, consistent with the conclusion that SHM in X-HIgM reflects an extra- or pre-GC process characterized by relatively decreased activity of AICDA and UNG compared to that characteristic of GC reactions.

Materials and methods

Tissue Sources

PB was obtained from a male X-HIgM patient at ages 9 and 13 who has two mutations, one of which introduces a stop codon in the CD40 ligand gene[40] and whose repertoire was previously analyzed at the age of 4 years.[5] Other X-HIgM patients included a 7-, 8-and 12-year-old and

three adults approximately 20 years old. Genetic phenotyping and clinical manifestations of these patients have been reported.[41] Apheresis of normal healthy donors, and six tonsil specimens from children aged 2-10 years were obtained with NIH IRB approval and informed consent in accordance with the Declaration of Helsinki.

Tissue processing

PBMCs were isolated from PB or splenic cells diluted in PBS using Ficoll-Hypaque (Pharmacia, Piscataway, NJ) separation. For microarray analysis, B cells were negatively selected using an EasySep™ magnetic column (StemCell Technologies, Vancouver, Canada) following the manufacturer's instructions. Purity was over 95%. Tonsils were minced and digested in RPMI medium (Life Technologies, Grand Island, NY) containing 210 U/ml collagenase type I (Worthington Biochemical Corp., Lakewood, New Jersey, USA) and 90 U/ml DNase (Sigma-Aldrich) for 30 minutes at 37°C. Following filtration through a wire mesh, the cells were washed twice in 20% FBS/RPMI and once in 10% FBS-RPMI before centrifugation over diatrizoate/Ficoll gradients.

Flow cytometric separation and analysis of B cell subpopulations

Cells were stained with the following fluorochrome conjugated monoclonal antibodies: mouse IgG1 anti-human CD19 PE (Becton Dickinson), mouse IgG2 anti-human IgD FITC (Pharmingen) or goat anti-human IgD FITC (Caltag), anti-CD38 APC (BD Biosciences) and anti-CD27-PE and isotype controls (Becton Dickinson Pharmingen, San Diego, CA). 7-aminoactinomycin (7-AAD; Molecular Probes, Eugene, OR) was used to exclude dead cells. PB and splenic cells were sorted according to CD19 and IgD or CD19, IgD and CD27 expression,

respectively. CD19$^+$ tonsillar B cells were sorted into subsets of naïve CD38$^-$IgD$^+$, pre-GC

founder CD38^{++}IgD$^+$ and GC CD38^{++}IgD$^-$ cells. Sorting was performed on a Becton Dickinson

FACS Vantage SE DiVa or a MoFlo Cell Sorter (DAKO). Additional samples were collected

for analysis on a FACS Calibur (Becton Dickinson) or CyAN (DAKO-Cytomation, Fort Collins,

CO) and analyzed using FlowJo software (TreeStar, Stanford University, CA), Paint-a-Gate

(Becton Dickinson), CellQuest (Becton Dickinson) and Summit (DAKO-Cyomation).

Lysate preparation and single cell PCR

PCR of V_H genes obtained from genomic DNA of individual B cells was carried out as described

previously.[42] The PCR error established for this method is approximately 6 mutations per 75,000

base pairs analyzed (.008%).[43,44] This PCR error rate indicates that the maximal number of

PCR-induced base pair changes in the nonproductive X-HIgM repertoire was 3.

Database and Sequence analysis

Germline sequences of nonproductive and productive rearrangements were determined by using

the JOINSOLVER® sequence analysis program.[45] The normal donor (ND) rearrangements

analyzed included 71 nonproductive and 414 productive CD19$^+$ rearrangements (GenBank

numbers X87006-X87082[46] (one donor) and Z80363-Z80770 (two donors),[47] 20 nonproductive

and 128 productive IgD$^+$CD27$^+$ (EF542547-EF542687) and 12 nonproductive and 100

productive IgD$^-$CD27$^+$ (EF542688-EF542796) rearrangements (one donor). The CD19$^+$ data

from X-HIgM patients included 209 nonproductive and 926 productive rearrangements from a

patient at ages 4 (AF077410-AF077525)[5] and 9 years (EF542103-EF542275), three other X-

HIgM children at ages 7-12 (EF541488-EF542102) and 3 X-HIgM adults (EF542315-

EF542546). The IgD$^+$CD27$^+$ X-HIgM repertoire consisting of 7 nonproductive and 32

productive rearrangements (EF542276-EF542314) were from a 13 year-old patient who was also

analyzed at ages 4 and 9. From the 9 yo X-HIgM patient, 72 nonproductive and 53 productive

light chain rearrangements were analyzed (EU875398-EU875522).

Seven new polymorphisms within codons 25-94 in 6 VH genes were identified from the

X-HIgM B cells and were not included in the mutation analysis. The 16 germline V$_H$ genes

utilized in the mutated nonproductive X-HIgM repertoire were used to calculate the expected

R:S ratios from the codons within FRS4, 2 and 3 (R:S = 3.2) and the codons from CDR1 and 2

(R:S = 4.5).

No gross abnormalities were noted in the repertoire that could contribute to an abnormal

mutation pattern (unpublished data).

Statistical Analysis

Differences between the mutation pattern of X-HIgM patients and NDs were determined by

using the χ^2 test for 2x2 tables with and without Yates correction.   The chi squared test was

valid because the total number of mutations in the nonproductive (n=51) and productive (n=251)

repertoire was > 40[48,49].

The standard error was propagated for mutational targeting of the hypermutable

nucleotide inside versus outside a motif using the formula $s = \sqrt{\dfrac{P(1-P)}{n}}$ where P is the

proportion of mutated bases either inside or outside the hypermutable position and n is the

number of bases (either mutated or unmutated) at the hypermutable position.   The standard error

of the fold difference is given by the formula $s_{\frac{in}{out}} = \dfrac{P_{in}}{P_{out}} \sqrt{\left(\dfrac{s_{in}}{P_{in}}\right)^2 + \left(\dfrac{s_{out}}{P_{out}}\right)^2}$ where P$_{in}$ (P$_{out}$) is the

proportion of mutated bases inside (outside) the hypermutable position with standard error s_{in} (s_{out}). Hypothesis testing was performed to compare the fold difference between X-HIgM and normal donors and to determine whether the observed frequency of random mutations was significantly different from the expected frequency. A 2-sigma confidence interval, which corresponds to a p-value slightly smaller than 0.05 surrounding the mean was used as the non-rejection region. For all statistical analyses $p < 0.05$ limit was used to determine significant differences.

Microarray analysis of gene expression

Total RNA was isolated from B cells using TRIZOL (Invitrogen, Carlsbad, California), purified using RNeasy mini Kit (Qiagen, Valencia, CA), processed for microarray analysis using standard Affymetrix protocols (www.affymetrix.com) and hybridized to HG-U133A Affymetrix Genechips®. Genechips® were scanned on a high resolution Affymetrix scanner using GCOS version 1.2 software. In addition, naïve and germinal center cells from 3 additional tonsils were profiled for gene expression using the Affymetrix U133plus 2.0 array. To determine changes in gene expression, the microarray signals from tonsillar naïve B cells (CD38$^-$IgD$^+$) were compared to activated naïve/pre-GC founder cells (CD38^{++}IgD$^+$) and tonsillar GC cells (CD38^{++}IgD$^-$) signals. Data were normalized according to standard Affymetrix algorithm analysis (MAS5) by scaling the median expression to 500, using methods identical to those described previously [50]. The raw data were log$_2$ transformed and entered into a database. Array elements with median signal intensities of less than 6 log2 units across the samples –approximately 25% of array elements—were removed from the analysis entirely in order to exclude poorly measured genes and genes not appreciably expressed in a sample. The array data were also reanalyzed after

normalization with the RMA algorithm [51]. The data have been deposited in the publicly

available database GEO (Accession numbers: GSE12845 and GSE12366). For comparison of

genes in different subpopulations or genes affected by CD154 treatment a two-tailed Student's

T-test was applied to individual comparisons. Both a 1.4-fold change or greater and a *p* value

less than 0.05 were considered significant. Further analyses were performed using a hierarchical

clustering algorithm[52] to group genes according to similar expression levels.

Results

Mutation Frequency of X-HIgM B cells

Initially, the mutation frequency of V_H genes was assessed for comparison between children and

adults with X-HIgM. In a ND, 85% of the nonproductive Ig rearrangements in $CD19^+$ B cells

were mutated compared to only 10% in X-HIgM ages 7-12 years, 1% in the 9 yo X-HIgM and

4% in adult X-HIgM patients (Table 1). Although the 4-year-old X-HIgM patient had more

mutated rearrangements (6/23; 26%) than all other X-HIgM patients, the frequency of mutated

sequences five years later had fallen closer to the level of others in his age group and the adults.

The overall mutation frequency of nonproductive rearrangements from all X-HIgM patients was

0.2 + 0.03% (mean + SEM) compared to 3.8% in the ND $CD19^+$ cells. Even when only the

mutated non-productive sequences were considered, there was a significant reduction in the X-

HIgM versus normal (0.9% + 0.4 versus 4.5%; $p<.0001$). Although the mutation frequency is

low, it is significantly greater than the PCR error (0.008%; $p<.0001$) established for this

method.[43] In the productive repertoire, adult patients had the highest overall mutation frequency

among the X-HIgM patients. However, it was more than 10-fold less (0.2% versus 2.5%) than

the overall mutation frequency of normal $CD19^+$ cells. Moreover, the mutational frequency of

just those sequences with mutated V_H rearrangements was almost 3-fold (1.2% versus 3.3%) less than the mutation frequency of normal adult mutated sequences. As opposed to NDs, the majority of mutated X-HIgM sequences contained only a single mutation (Figure 1). Similar results were observed when light chain mutations were analyzed (Table S1). Of the 15,872 base pairs analyzed in nonproductive light chain rearrangements, there were 26 mutations for an overall mutational frequency of 0.2%. This is significantly ($p<.001$ with Yates correction) lower than the mutational frequency of normal light chains as previously reported[53,54].

The size of $IgM^+IgD^+CD27^+$ memory B cell population and its origin in X-HIgM patients is controversial.[6-8] We, therefore, analyzed the frequency of $CD27^+$ memory B cells in a ND and four X-HIgM patients of various ages. In these X-HIgM patients, less than 3% of the $CD19^+$ population was $CD27^+$, significantly ($p<0.0001$ with Yates correction) less than the $CD27^+$ population in NDs (Figure 2). Notably, the post-switch IgD^-CD27^+ population was the most markedly deficient (26% in ND versus 0.6% in X-HIgM). To examine the mutational status of $CD27^+$ cells in X-HIgM, we sequenced 78 V_H rearrangements obtained from IgD^+CD27^+ cells of a 13 yo X-HIgM patient previously analyzed at ages 4 and 9 years, and compared them to V_H rearrangements from 148 pre-switched (IgD^+CD27^+) memory B cells of a normal young adult. In normals, 90-98% of the nonproductive and productive $CD27^+IgD^+$ pre-switched memory cells expressed mutated V_H genes (Table S2).[55] In X-HIgM, only 1/7 IgD^+CD27^+ B cells expressed mutated V_H genes in the nonproductive repertoire and in the productive repertoire 12/32 (38%) were mutated (data not shown). The overall mutation frequency of nonproductive (13/1496, 0.9%) and productive (73/7809, 0.9%) rearrangements from X-HIgM IgD^+CD27^+ B cells was significantly less than the mutation frequency of rearrangements from normal IgD^+CD27^+ memory B cells (160/3694, 4.3%; $p<0.001$). As a result, the IgD^+CD27^+ and IgD^+CD19^+ X-

HIgM repertoires were dominated by sequences with 0-5 mutations, whereas the ND IgD^+CD27^+

repertoire was enriched in sequences with 5 or more mutations per sequence. Since the X-HIgM

IgD^+CD27^+ rearrangements were similar in mutation frequency to the X-HIgM mutated

IgD^+CD19^+ rearrangements and recent evidence has identified a $CD27^-$ human memory B cell

population,[56-58] the X-HIgM data were combined in order to compare all mutations to similarly

pooled ND B cell subsets (Table S3). It is notable that the mutation frequency and the targeting

of mutations to specific motifs (see below) were very similar in X-HIgM nonproductive and

productive repertoires, indicating that there was minimal selection of these mutations.

Mutation focusing in hot spots

In X-HIgM, RGYW and WRCY motifs were mutated. No significant differences were found

between the percentage of mutations in the nonproductive X-HIgM RGYW and WRCY motifs

in comparison to that expected from random chance (12.5%) (Table 2) and no significant

differences were found (p=0.16) between X-HIgM and ND nonproductive or productive

repertoires. A and T mutations in WA and TW motifs, reflecting the impact of POL η activity,

were also similar in X-HIgM and NDs. Consistent with the preferential targeting of WA motifs

by POL η,[59] the mutations in TA motifs significantly exceeded that expected from random

chance (6.35%) and other dinucleotides of WA/TW in NDs (71.4/593; 12%, $p<.0001$) and

although not significantly greater, a trend for an increase in TA mutations occurred in X-HIgM

(data not shown). Notably, 64% of the mutations in the nonproductive X-HIgM repertoire could

be accounted for by mutations in RGYW/WRCY or WA/TW compared to 73% in NDs.

Some differences between X-HIgM and NDs became apparent by examining the ratio of

replacement (R) to silent (S) mutations. Notably, the R:S ratios in the CDRs and FRs of

nonproductive X-HIgM rearrangements were approximately half as great as in NDs. Unlike

NDs, no significant difference was found between the FR and CDR R:S ratios (Table 2).

Finally, in contrast to NDs in which R mutations in the CDRs are eliminated from the productive

repertoire (p=.013; with Yates correction p=0.17)[60], no significant difference was found in the

CDR R:S ratios between the nonproductive and productive repertoire of X-HIgM B cells.

Transition substitutions dominate in X-HIgM

X-HIgM B cells also have an abnormal ratio of transitions:transversions. Although the ratio of

transition:transversion substitution would be 1:2 if these occurred randomly, transitions were

nearly as common as transversions in normal B cell SHM (Table 2). In contrast, in the

nonproductive X-HIgM repertoire, transitions were favored (33:18 or 1.8:1). A similar finding

was noted with light chain mutations in which transitions dominated (18:8 or 2.3) (Table S1).

When both heavy and light chain mutations were considered there was a significant difference in

their distribution in X-HIgM versus normal (51 transitions:26 transversions in X-HIgM versus

499 transitions:484 transversions in normal; p=0.008 (p=0.01 with Yates correction). Transitions

tended to be negatively selected in X-HIgM VH genes and as a result the ratio of

transitions:transversions in the productive X-HIgM repertoire was closer to normal (1.3:1).

To examine the increased frequency of transitions over transversions in X-HIgM in more

detail, point mutations of individual nucleotides in X-HIgM rearrangements were analyzed. The

increase in total transitions in X-HIgM versus normal nonproductive rearrangements were biased

toward an increased presence of C→T transitions (Figure 3A). (In X-HIgM, 13 of the 51

mutations (25%) were C→T substitutions compared to 60/593 (10%) in NDs.) Within the

population of C mutations in the HC repertoire, significantly (p=.005 or .01 with Yates

correction) more transitions than transversions were found in X-HIgM compared to normal donors (13 C>T/2 C transversions versus 60 C>T/63 transversions, respectively). A significant difference ($p=.04$) was also noted in the productive repertoire in which 36/49 C mutations (73%) were C→T transitions (Figure 3B).

Targeting the hypermutable G/C in RGYW/WRCY motifs

Since AICDA is known to target the C in WRCY motifs, we examined the specific targeting of G/C in RGYW/WRCY motifs as a fingerprint of AICDA activity. Initially, the distribution of individual nucleotide mutations was analyzed in the nonproductive repertoire of X-HIgM. Mutations in all four nucleotides were comparable (Figure 3A). In contrast, NDs mutate A or G > C > T and exhibit significant G versus C ($p<.0001$) and A versus T ($p<.0001$) strand bias in both nonproductive and productive repertoires. In X-HIgM, there is a bias favoring A mutations more than T mutations ($p<.0001$) and G mutations more than C mutations ($p<.0001$) only in the productive repertoires, presumably as a consequence of selection (Figure 3B).

Even though the percentage of mutations in RGYW/WRCY motifs was comparable to NDs, the abnormal nucleotide substitution pattern prompted an analysis of mutations in specific residues in these motifs to assess the relative activity of AICDA in X-HIgM. The mutation frequency of the hypermutable nucleotide (C or G) in the targeted motifs (WRCY or RGYW, respectively) was compared with the mutation frequency of the same nucleotide occurring elsewhere in the variable segment. Marked differences were noted when an analysis of G mutations in the second position of RGYW motifs was carried out (Table 3). In normal B cells, the mutation frequency of G nucleotides in RGYW was 5.5-6.5-fold greater than those outside RGYW in productive and nonproductive repertoires, respectively. In contrast, in the

nonproductive X-HIgM repertoire, targeted mutation of the hypermutable G in RGYW motifs was absent since the frequency of G mutation inside RGYW was no greater than the frequency of mutation outside RGYW (0.11% and 0.09%; respectively). The possibility that the difference in the distribution of G mutation between X-HIgM and normal donors was related to random chance was highly unlikely because the fold difference comparing G mutations in and out of RGYW motifs in X-HIgM versus normal donors was 1.25 ± 0.97 versus 6.48 ± 0.9, respectively ($p<0.0001$). Even though there was evidence of selection of G mutations within RGYW in the X-HIgM productive repertoire, the fold increase in G mutations within RGYW compared to those outside RGYW remained significantly less than in the normal productive repertoire (2.7 versus 5.4, $p<.0001$). A loss of hypermutable G mutations in RGYW motifs was also noted in the nonproductive light chains (Table S1).

In contrast, pronounced targeting of C mutation in WRCY was found in X-HIgM as well as in normals in the nonproductive and productive repertoires. In normal B cells, the difference in the mutation frequency of C inside WRCY versus all other C mutations was 4.4-fold in nonproductive rearrangements. Similarly, of the 15 C mutations in X-HIgM, half occurred in the hypermutable position of WRCY resulting in a 7-fold greater C mutational frequency inside WRCY versus all other C mutations (0.5% versus 0.07%; $p<.0001$ with or without Yates correction). In the nonproductive light chain rearrangements, there was greater than 3-fold more hypermutable C targeting in WRCY compared to all other C mutations (Table S1).

Analysis of mutations in the other three positions of R–YW motifs in the nonproductive X-HIgM repertoire, revealed a 3.6-fold increase in the frequency of A mutations inside RGYW (0.25%) compared to outside RGYW (0.07%), most of which were in overlapping WA/TW motifs. In sharp contrast, A mutations inside and outside WRCY were equal. A similar pattern

was noted in normals. Although WA mutations were found preferentially in RGYW motifs, there were numerous WA mutations outside RGYW/WRCY.

Gene expression profiling of GC B cells

The multiple mutation abnormalities in X-HIgM B cells, which include a decreased frequency of SHM, decreased frequency of G mutation in RGYW and decreased transversion substitutions, suggested CD40-CD154 signaling may regulate expression of genes involved in mutational targeting or DNA repair. To address this question, gene expression profiles were analyzed from tonsil B cell subsets including naive B cells (CD19$^+$CD38$^-$IgD$^+$), pre-GC B cells (CD19$^+$CD38^{++}IgD$^+$), and germinal center B cells (CD19$^+$CD38^{++}IgD$^-$), as well as PB naïve (CD19$^+$CD27$^+$) or memory (CD19$^+$CD27$^-$) B cells (Table 4 and Figs. S1, S2 and S3)

Initially a list of 170 genes known to contribute to MMR, base excision repair (BER), nucleotide excision repair (NER), TLS, DNA repair, recombination, switching, and nonhomologous end joining (NHEJ), transcription and DNA replication were investigated (Table S4). Among the genes examined, *aicda, rpa, exo 1, msh2, msh6, pcna, pol θ* and *ung* were expressed at increased levels in tonsillar pre-GC (CD19$^+$CD38^{++}IgD$^+$) and/or GC cells (CD19$^+$CD38^{++}IgD$^-$) compared to tonsil naïve B cells (CD19$^+$CD38$^-$IgD$^+$), similar to what has been noted by others [61]. Whereas *ung* and *ogg1* glycosylases were increased in normal GC B cells single-strand-selective monofunctional *ung*1 (*smug*1) was not increased in normal GC B cells. Genes involved in NER (*xpa, xpc, xpd, ercc*1 and *ercc*4) and NHEJ (*ku* 70, *ku* 80, and *ligase* 4) were also not increased but *xrcc*4 was significantly increased in pre-GC B cells. Whereas expression of rad50 and *rad51* were upregulated, other genes involved in homologous recombination, radiation sensitivity and double strand breaks did not change. Although multiple

polymerases expressed in B cells have been investigated to determine whether they contribute to SHM, only *pol* θ expression was increased (2-4-fold) in tonsillar GC cells compared to tonsil naïve B cells. Notably, in vitro stimulation of PB and splenic B cells with recombinant CD154 did not consistently alter expression of genes potentially involved in SHM (data not shown). These patterns of gene expression in tonsillar subsets were confirmed when a separate set of microarray data was acquired from three additional tonsil samples from normal donors (Figure S4).

Discussion

The current mutation analysis of individual B cells from X-HIgM patients has enabled a comparison with the mutation patterns in normals and characterization of the molecular mechanisms of SHM operating in the absence of CD40-CD154 interactions. The significantly lower mutation frequency accompanied by few CD27[+] memory B cells was sustained into adulthood in X-HIgM patients, confirming that the mutation defect resulting from the lack of CD40 signals is persistent and that CD40-CD154 interaction is an important step for both generation of memory cells and mutation amplification. Other defects found in X-HIgM included negligible class–switching, as noted previously,[62] and a much lower than normal frequency of CD27[+] B cells.[8] The current data reveal that CD40-CD154 signals had a greater impact on switching than on the frequency of CD27[+] B cells, although the persistent defect in both was notable, suggesting there is a hierarchical dependence on CD40 signals. Whereas CD27, a tumor necrosis factor (TNF) receptor family member, is noted as a memory B cell marker in humans,[63] less than 14% of the X-HIgM CD27[+] B cells were mutated. These data suggest that, in this circumstance, CD27 did not correlate with memory, as previously noted in AICDA deficient B

cells.[64] Importantly, the pattern of mutations in X-HIgM B cells is significantly different than that noted in normal CD27[+] memory B cells with no targeting of G in RGYW and significantly fewer transversions. The data are most consistent with the conclusion that X-HIgM B cells have experienced a diminished impact of AICDA and UNG presumably because they have not been exposed to GC reactions. The significantly different mutation frequencies and patterns in X-HIgM B cells versus normal CD27[+] IgD[+] pre-switch or CD27[+]IgD[-] post-switch memory B cells provides no support for the contention that the mutated X-HIgM B cells represent either of the normal memory B cell subsets as previously proposed.[7] Rather, these cells appear to be a population of pre-GC B cells that have experienced limited AICDA and UNG activity, presumably because they have not been exposed to GC reactions.

However, mutations clearly occurred in X-HIgM, with some features of normal SHM, including mutational targeting in WRCY and WA motifs, and, therefore, supported previous findings that not all mutation is solely dependent on CD40-CD154 stimulation and GC reactions.[5,6] In this regard, it is known that mutations occur in the shark and frog that lack GCs[65] and GCs are not required for mutational targeting of RGYW/WRCY motifs in response to immunization in the lymphotoxin α knockout mouse that lack GCs.[66] These findings are consistent with the conclusion that, in humans, SHM may be initiated before a GC reaction when the B cell encounters antigen and antigen activated T cells. However, the data also clearly demonstrate that most mutational activity or the survival of memory B cells with heavily mutated V_H genes requires CD40-CD154 signals[67] or GC reactions. In this regard, GC B cells exhibit increased expression of *aicda*, as previously shown,[68] as well as *ung* and selected DNA repair genes to achieve diversity and high frequency mutation.

One of the most striking features of the X-HIgM mutation pattern was the absence of G mutations in RGYW motifs. An important consequence of the loss of RGYW mutations was the decrease of R substitutions. This defect was most likely the consequence of 5-15-fold lower *aicda* expression in naive B cells compared to pre-GC and GC cells since a relationship between the level of expression of *aicda* and SHM has been reported.[69] However, *aicda* expression was apparently sufficient to initiate a pattern of mutations that provided insight into the basic mechanisms of SHM.

A second major abnormality in X-HIgM B cells was a marked decrease in transversions that presumably reflected decreased expression and/or availability of UNG2, another gene upregulated in GC B cells. A similar skewing toward transitions was attributed to the absence of UNG in a chicken cell line,[70] mouse[29] and human B cells[71] suggesting that UNG2 is less active in X-HIgM B cells and can explain the relatively increased frequency of transitions in X-HIgM. With the unavailability of UNG2 in X-HIgM, the AICDA-induced lesions were apparently copied over by a high fidelity polymerase, such as POL δ[39], that inserted thymidine in place of uracil during DNA synthesis generating C$\rightarrow$T transitions**.** In the absence of UNG, SMUG1 can remove uracils from DNA and therefore, SMUG1 is capable of partially complementing UNG2 in UNG deficient human cells in vitro[72] but not in human B cell lines.[73] These findings, along with our gene expression profiling that found downregulation of SMUG1 in GC cells, as previously noted in murine B cells,[74] are consistent with the conclusion that UNG is the major mechanism by which uracil is removed from DNA in human B cells and that the predominance of transitions in X-HIgM B cells relates to the relative insufficiency of UNG in the absence of GC reactions.

AICDA is known to target both strands of DNA,[27-29,75] thus, we concluded that the prominent targeting of C mutations in WRCY motifs, which is likely to occur as a direct deoxycytidine attack by AICDA on the transcribed strand, was an indication that AICDA was active in X-HIgM B cells. Although the overall mutation frequency in X-HIgM was considerably less than normal, it is notable that the frequency of C mutations in WRCY was nearly normal, implying that this specific mutational targeting did not depend on CD40-CD154 interactions or GC reactions. Importantly, these mutations are completely absent in subjects who lack expression of AICDA [76]. It was of particular interest to note, that although the percentage of all mutations that accumulated in RGYW motifs in X-HIgM was nearly comparable to that found in normal B cells, there were no preferential mutations in the most frequently mutated G position (RGYW). It was possible that decreased G mutations in RGYW might represent not only decreased AICDA activity, but also frequent repair of these AICDA-induced mutations. The mechanism responsible for repair of these mutations most likely involves mismatch repair enzymes associated with SHM,[39] at least some of which, such as MSH2/MSH6 and EXO1, might be available in B cells in the absence of GC reactions.

Even though the hypermutable G in RGYW was not mutated, the percentage of mutations in RGYW motifs in X-HIgM was nearly normal implying that mutations must have occurred in the other three positions (R_YW). The first and last nucleotides of RGYW motifs are frequently A residues that coincide with WA motifs and these A residues contributed to the mutations in RGYW in X-HIgM. Assuming that WA mutations result from the activity of POL η,[59] this mechanism is the most likely source of these mutations. Thus, the current data suggest that in X-HIgM there is preferential targeting of POL η to RGYW motifs, presumably because of the activity of AICDA. Even though the initial AICDA lesion is apparently repaired, the sequences retain the fingerprint of

POL η activity in RGYW sequences. A similar fingerprint could be seen in RGYW motifs in NDs, implying that preferential targeting of POL η to RGYW motifs was a general feature of SHM independent of GC reactions. It is notable that induction of A mutations has recently been demonstrated when a single C is introduced into the nontranscribed strand of a transgenic SHM substrate by recruiting mismatch repair that presumably includes POL η[75] and establishes a linkage between AICDA and POL η activities.

The presence of WA mutations in X-HIgM B cells suggested that POL η is active in the absence of GC reactions. However, the frequency of POL η related mutations was much greater in normal B cells, even though expression of *pol η* was not increased in GC cells. It has recently been noted that some POL η activity might require the activity of MSH2.[33] Of note, MSH2 was upregulated 1.6-2.1-fold in GC cells. Thus, the apparent increased activity of POL η in normal B cells may relate to the upregulation of MSH2 and its capacity to recruit POL η to the site of mutation. This conclusion is consistent with the recent observation that the discrepancy in the induction of A mutations when a C is introduced in the transcribed strand is overcome when MSH2 is absent[75] and is compatible with a role for MSH2 in recruitment of POL η to the targeted motifs during SHM.

In summary, the X-HIgM B cell repertoire has few CD27[+] memory B cells and the quantity and pattern of mutation suggests a low level of AICDA activity initiates SHM but a GC reaction is required for amplification and upregulation of genes that are required for the full spectrum of mutation.

Authorship

Contribution: N.S.L., P.L.L., A.C.G. and P.E.L. designed the research protocol, N.S.L., S.Y. and P.H.L.K. prepared the sequence data. D.E.R. performed statistical analyses, W.Z. performed the CD154 in vitro stimulation, P.L.L., D.J.J. and S.S.D. performed the microarray analysis and N.S.L. did the mutation analysis and wrote the manuscript. N.S.L. and P.E.L. contributed to the final preparation of the manuscript.

Conflict of interest disclosure: The authors have no competing financial interests.

References

1.   DiSanto JP, Bonnefoy JY, Gauchat JF, Fischer A, de Saint Basile G. CD40 ligand mutations in x-linked immunodeficiency with hyper-IgM. 0028-0836. 1993;361:541-543.

2.   Allen RC, Armitage RJ, Conley ME, et al. CD40 ligand gene defects responsible for X-linked hyper-IgM syndrome. Science. 1993;259:990-993.

3.   Longo NS, Lipsky PE. Why do B cells mutate their immunoglobulin receptors? Trends Immunol. 2006;27:374-380.

4.   Kroczek RA, Graf D, Brugnoni D, et al. Defective expression of CD40 ligand on T cells causes "X-linked immunodeficiency with hyper-IgM (HIGM1)". Immunol Rev. 1994;138:39-59.

5.   Brezinschek HP, Dörner T, Monson NL, Brezinschek RI, Lipsky PE. The influence of CD40-CD154 interactions on the expressed human V(H) repertoire: analysis of V(H) genes expressed by individual B cells of a patient with X-linked hyper-IgM syndrome. Int Immunol. 2000;12:767-775.

6.   Weller S, Faili A, Garcia C, et al. CD40-CD40L independent Ig gene hypermutation suggests a second B cell diversification pathway in humans. Proc Natl Acad Sci U S A. 2001;98:1166-1170.

7.   Weller S, Braun MC, Tan BK, et al. Human blood IgM "memory" B cells are circulating splenic marginal zone B cells harboring a prediversified immunoglobulin repertoire. Blood. 2004;104:3647-3654.

8.   Agematsu K, Nagumo H, Shinozaki K, et al. Absence of IgD-CD27(+) memory B cell population in X-linked hyper-IgM syndrome. J Clin Invest. 1998;102:853-860.

9. Jacob J, Kelsoe G. In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. II. A common clonal origin for periarteriolar lymphoid sheath-associated foci and germinal centers. J Exp Med. 1992;176:679-687.

10. Leanderson T, Kallberg E, Gray D. Expansion, selection and mutation of antigen-specific B cells in germinal centers. Immunol Rev. 1992;126:47-61.

11. Dörner T, Foster SJ, Brezinschek HP, Lipsky PE. Analysis of the targeting of the hypermutational machinery and the impact of subsequent selection on the distribution of nucleotide changes in human VHDJH rearrangements. Immunol Rev. 1998;162:161-171.

12. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. 0092-8674. 2000;102:553-563.

13. Petersen-Mahrt SK, Harris RS, Neuberger MS. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. 0028-0836. 2002;418:99-103.

14. He B, Qiao X, Cerutti A. CpG DNA induces IgG class switch DNA recombination by activating human B cells through an innate pathway that requires TLR9 and cooperates with IL-10. J Immunol. 2004;173:4479-4491.

15. Gourzi P, Leonova T, Papavasiliou FN. Viral induction of AID is independent of the interferon and the Toll-like receptor signaling pathways but requires NF-kappaB. J Exp Med. 2007;204:259-265.

16. Ettinger R, Sims GP, Fairhurst AM, et al. IL-21 Induces Differentiation of Human Naive and Memory B Cells into Antibody-Secreting Plasma Cells. J Immunol. 2005;175:7867-7879.

17.   Moldenhauer G, Popov SW, Wotschke B, et al. AID expression identifies interfollicular large B cells as putative precursors of mature B-cell malignancies. Blood. 2006;107:2470-2473.

18.   Cattoretti G, Buttner M, Shaknovich R, Kremmer E, Alobeid B, Niedobitek G. Nuclear and cytoplasmic AID in extrafollicular and germinal center B cells. Blood. 2006;107:3967-3975.

19.   Bergqvist P, Gardby E, Stensson A, Bemark M, Lycke NY. Gut IgA class switch recombination in the absence of CD40 does not occur in the lamina propria and is independent of germinal centers. J Immunol. 2006;177:7772-7783.

20.   He B, Xu W, Santini PA, et al. Intestinal bacteria trigger T cell-independent immunoglobulin A(2) class switching by inducing epithelial-cell secretion of the cytokine APRIL. 1074-7613. 2007;26:812-826.

21.   Willenbrock K, Jungnickel B, Hansmann ML, Kuppers R. Human splenic marginal zone B cells lack expression of activation-induced cytidine deaminase. Eur J Immunol. 2005;35:3002-3007.

22.   Tangye SG, Good KL. Human IgM+CD27+ B Cells: Memory B Cells or "Memory" B Cells? J Immunol. 2007;179:13-19.

23.   Rogozin IB, Kolchanov NA. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. Biochim Biophys Acta. 1992;1171:11-18.

24.   Dörner T, Foster SJ, Farner NL, Lipsky PE. Somatic hypermutation of human immunoglobulin heavy chain genes: targeting of RGYW motifs on both DNA strands. Eur J Immunol. 1998;28:3384-3396.

25. Ramiro AR, Stavropoulos P, Jankovic M, Nussenzweig MC. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. Nat Immunol. 2003;4:452-456.

26. Peters A, Storb U. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. 1074-7613. 1996;4:57-65.

27. Shen HM, Storb U. Activation-induced cytidine deaminase (AID) can target both DNA strands when the DNA is supercoiled. Proc Natl Acad Sci U S A. 2004;101:12997-13002.

28. Rada C, Ehrenstein MR, Neuberger MS, Milstein C. Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. 1074-7613. 1998;9:135-141.

29. Rada C, Williams GT, Nilsen H, Barnes DE, Lindahl T, Neuberger MS. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. Curr Biol. 2002;12:1748-1755.

30. Phung QH, Winter DB, Cranston A, et al. Increased hypermutation at G and C nucleotides in immunoglobulin variable genes from mice deficient in the MSH2 mismatch repair protein. J Exp Med. 1998;187:1745-1751.

31. Wiesendanger M, Kneitz B, Edelmann W, Scharff MD. Somatic hypermutation in MutS homologue (MSH)3-, MSH6-, and MSH3/MSH6-deficient mice reveals a role for the MSH2-MSH6 heterodimer in modulating the base substitution pattern. J Exp Med. 2000;191:579-584.

32. Martomo SA, Yang WW, Gearhart PJ. A role for Msh6 but not Msh3 in somatic hypermutation and class switch recombination. J Exp Med. 2004;200:61-68.

33.   Delbos F, Aoufouchi S, Faili A, Weill JC, Reynaud CA. DNA polymerase {eta} is the sole contributor of A/T modifications during immunoglobulin gene hypermutation in the mouse. J Exp Med. 2007;204:17-23.

34.   Bardwell PD, Woo CJ, Wei K, et al. Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice. Nat Immunol. 2004;5:224-229.

35.   Wilson TM, Vaisman A, Martomo SA, et al. MSH2-MSH6 stimulates DNA polymerase eta, suggesting a role for A:T mutations in antibody genes. J Exp Med. 2005;201:637-645.

36.   Martomo SA, Yang WW, Wersto RP, et al. Different mutation signatures in DNA polymerase eta- and MSH6-deficient mice suggest separate roles in antibody diversification. Proc Natl Acad Sci U S A. 2005;102:8656-8661.

37.   Yavuz S, Yavuz AS, Kraemer KH, Lipsky PE. The role of polymerase eta in somatic hypermutation determined by analysis of mutations in a patient with xeroderma pigmentosum variant. J Immunol. 2002;169:3825-3830.

38.   Zeng X, Negrete GA, Kasmer C, Yang WW, Gearhart PJ. Absence of DNA polymerase eta reveals targeting of C mutations on the nontranscribed strand in immunoglobulin switch regions. J Exp Med. 2004;199:917-924.

39.   Casali P, Pal Z, Xu Z, Zan H. DNA repair in antibody somatic hypermutation. Trends Immunol. 2006;27:313-321.

40.   Grammer AC, Bergman, M. C., Miura, Y., Fujita, K., Davis, L.S. and Lipsky, P. E. The CD40 ligand expressed by human B cells costimulates B cell responses. 0022-1767. 1995;154:4996.

41.   Jain A, Atkinson TP, Lipsky PE, Slater JE, Nelson DL, Strober W. Defects of T-cell effector function and post-thymic maturation in X-linked hyper-IgM syndrome. J Clin Invest. 1999;103:1151-1158.

42.   Brezinschek HP, Foster SJ, Brezinschek RI, Dörner T, Domiati-Saad R, Lipsky PE. Analysis of the human VH gene repertoire. Differential effects of selection and somatic hypermutation on human peripheral CD5(+)/IgM+ and CD5(-)/IgM+ B cells. J Clin Invest. 1997;99:2488-2501.

43.   Yavuz AS, Lipsky PE, Yavuz S, Metcalfe DD, Akin C. Evidence for the involvement of a hematopoietic progenitor cell in systemic mastocytosis from single-cell analysis of mutations in the c-kit gene. Blood. 2002;100:661-665.

44.   Farner NL, Dörner T, Lipsky PE. Molecular mechanisms and selection influence the generation of the human V lambda J lambda repertoire. J Immunol. 1999;162:2137-2145.

45.   Souto-Carneiro MM, Longo NS, Russ DE, Sun H-w, Lipsky PE. Characterization of the human Ig heavy chain antigen binding complementarity determining region 3 using a newly developed software algorithm, JOINSOLVER. J Immunol. 2004;172:6790-6802.

46.   Brezinschek HP, Dörner T, Monson N, Brezinschek RI, Lipsky PE. The influence of CD40-CD154 interactions on the expressed human VH repertoire: analysis of VH genes expressed by individual B cells of a patient with X-linked hyper-IgM syndrome. Int Immunol. 2000;12:767-775.

47.   Dörner T, Brezinschek HP, Foster SJ, Brezinschek RI, Farner NL, Lipsky PE. Delineation of selective influences shaping the mutated expressed human Ig heavy chain repertoire. J Immunol. 1998;160:2831-2841.

48.   Haviland MG. Yates's correction for continuity and the analysis of 2 x 2 contingency tables. Stat Med. 1990;9:363-367; discussion 369-383.

49.   Cochran WG. Some Methods for Strengthening the Common χ2 Tests. 0006-341X. 1954;10:417-451.

50.   Dave SS, Fu K, Wright GW, et al. Molecular diagnosis of Burkitt's lymphoma. N Engl J Med. 2006;354:2431-2442.

51.   Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics. 2003;4:249-264.

52.   Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A. 1998;95:14863-14868.

53.   Foster SJ, Dörner T, Lipsky PE. Somatic hypermutation of VkappaJkappa rearrangements: targeting of RGYW motifs on both DNA strands and preferential selection of mutated codons within RGYW motifs. Eur J Immunol. 1999;29:4011-4021.

54.   Monson NL, Dorner T, Lipsky PE. Targeting and selection of mutations in human Vlambda rearrangements. Eur J Immunol. 2000;30:1597-1605.

55.   Tian C, Luskin GK, Dischert KM, Higginbotham JN, Shepherd BE, Crowe JE, Jr. Evidence for preferential Ig gene usage and differential TdT and exonuclease activities in human naive and memory B cells. Mol Immunol. 2007;44:2173-2183.

56.   Wei C, Anolik J, Cappione A, et al. A new population of cells lacking expression of CD27 represents a notable component of the B cell memory compartment in systemic lupus erythematosus. J Immunol. 2007;178:6624-6633.

57.   Wirths S, Lanzavecchia A. ABCB1 transporter discriminates human resting naive B cells from cycling transitional and memory B cells. Eur J Immunol. 2005;35:3433-3441.

58.   Ehrhardt GR, Hsu JT, Gartland L, et al. Expression of the immunoregulatory molecule FcRH4 defines a distinctive tissue-based population of memory B cells. J Exp Med. 2005;202:783-791.

59.   Rogozin IB, Pavlov YI, Bebenek K, Matsuda T, Kunkel TA. Somatic mutation hotspots correlate with DNA polymerase eta error spectrum. Nat Immunol. 2001;2:530-536.

60.   Dörner T, Brezinschek HP, Brezinschek RI, Foster SJ, Domiati-Saad R, Lipsky PE. Analysis of the frequency and pattern of somatic mutations within nonproductively rearranged human variable heavy chain genes. J Immunol. 1997;158:2779-2789.

61.   Hogerkorp CM, Borrebaeck CA. The human CD77- B cell population represents a heterogeneous subset of cells comprising centroblasts, centrocytes, and plasmablasts, prompting phenotypical revision. J Immunol. 2006;177:4341-4349.

62.   Notarangelo LD, Duse M, Ugazio AG. Immunodeficiency with hyper-IgM (HIM). Immunodefic Rev. 1992;3:101-121.

63.   Agematsu K, Hokibara S, Nagumo H, Komiyama A. CD27: a memory B-cell marker. Immunol Today. 2000;21:204-206.

64.   Revy P, Muto T, Levy Y, et al. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). 0092-8674. 2000;102:565-575.

65.   Diaz M, Flajnik MF. Evolution of somatic hypermutation and gene conversion in adaptive immunity. Immunol Rev. 1998;162:13-24.

66.   Matsumoto M, Lo SF, Carruthers CJ, et al. Affinity maturation without germinal centres in lymphotoxin-alpha-deficient mice. 0028-0836. 1996;382:462-466.

67.   Foy TM, Laman JD, Ledbetter JA, Aruffo A, Claassen E, Noelle RJ. gp39-CD40 interactions are essential for germinal center formation and the development of B cell memory. J Exp Med. 1994;180:157-163.

68.   Muramatsu M, Sankaranand VS, Anant S, et al. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. J Biol Chem. 1999;274:18470-18476.

69.   Ruckerl F, Bachl J. Activation-induced cytidine deaminase fails to induce a mutator phenotype in the human pre-B cell line Nalm-6. Eur J Immunol. 2005;35:290-298.

70.   Di Noia J, Neuberger MS. Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. 0028-0836. 2002;419:43-48.

71.   Imai K, Slupphaug G, Lee WI, et al. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. Nat Immunol. 2003;4:1023-1028.

72.   Akbari M, Otterlei M, Pena-Diaz J, et al. Repair of U/G and U/A in DNA by UNG2-associated repair complexes takes place predominantly by short-patch repair both in proliferating and growth-arrested cells. Nucleic Acids Res. 2004;32:5486-5498.

73.   Kavli B, Andersen S, Otterlei M, et al. B cells from hyper-IgM patients carrying UNG mutations lack ability to remove uracil from ssDNA and have elevated genomic uracil. J Exp Med. 2005;201:2011-2021.

74.   Di Noia JM, Rada C, Neuberger MS. SMUG1 is able to excise uracil from immunoglobulin genes: insight into mutation versus repair. Embo J. 2006;25:585-595.

75.   Unniraman S, Schatz DG. Strand-biased spreading of mutations during somatic hypermutation. Science. 2007;317:1227-1230.

76.	Longo NS, Satorius CL, Plebani A, Durandy A, Lipsky PE. Characterization of Ig gene somatic hypermutation in the absence of activation-induced cytidine deaminase. J Immunol. 2008;181:1299-1306.

Figure 1. Nonproductive and productive repertoires in X-HIgM are significantly less mutated than NDs.   All mutations in X-HIgM and normal donors were grouped according to the number of mutations/mutated sequence. (dashed line X-HIgM; solid line NDs) A. Nonproductive repertoire, B. Productive repertoire.

Figure 2. CD27[+] B cells are decreased in X-HIgM. PBMCs from X-HIgM patients and NDs were stained with CD19, IgD and CD27 fluorochrome conjugated antibodies and analyzed by flow cytometry using CELLQuest software™. The percentage of naïve (CD19[+]IgD[+]CD27[-]), pre-switch (CD19[+]IgD[+]CD27[+]) and post-switch (CD19[+]IgD[-]CD27[+]) cells represent the data analyzed from four 7-13 year old X-HIgM patients and four NDs. *Significant ($p<0.001$) difference in the percentage of cells between ND and X-HIgM.

Figure 3. Cytosine to thymine transition substitutions are increased in X-HIgM B cells. The number (first row) and percentage (second row) of individual nucleotide substitutions detected in the variable segment of (A) nonproductive and (B) productive VDJH rearrangements in PB B cells from 4 normal donors and 7 X-HIgM patients are presented. The tables representing the percentages of all mutations scored are corrected for base composition of the target sequence. *Significant difference in C > T transitions; †significant difference in G versus C mutation frequency; ‡significant difference in A versus T mutation frequency.

Table 1. The mutation frequency of heavy chain rearrangements from IgD$^+$CD19$^+$ B cells of individual X-HIgM patients of different ages

	4 yo X-HIgM		9 yo X-HIgM		7-12 yo X-HIgM		Adult X-HIgM		ND Adult	
	NP	P	NP	P	NP	P	NP	P	NP	P
Total Number of Rearrangements	23	92	31	143	104	509	51	182	13	69
Percentage of Mutated Rearrangements	6/23 ✻ 26%	14/92 15%	3/331 1%	11/143 8%	10/104 10%	38/509 8%	2/51 4%	17/182 9%	11/13 ✻ 85%	53/69 77%
Total Number of Mutations	10	18	3	11	14	103	11	48	126	440
Overall Mutation Frequency	10/5814 † 0.2%	18/23208 0.08%	3/6540 0.05%	11/31315 0.04%	14/22461 0.06%	103/100669 0.1%	11/11345 0.1%	48/40687 0.2%	126/3330 † 3.8%	440/17579 2.5%
Mutation Frequency of Mutated Rearrangements	10/1524 0.7%	18/3540 0.6%	3/644 0.5%	11/2377 0.5%	14/2022 0.7%	103/7548 1.4%	11/465 ‡ 2.4%	48/4126 1.2%	126/2826 ‡ 4.5%	440/13484 3.3%

The data for the 4 yo and 9 yo are from the same patient after a 5 year interval. The data from three different patients were combined in each of the 7-12-year-old and adult X-HIgM categories. The normal donor data include GenBank accession numbers X87006-87089 (excluding X87017 and X87083). The mutation frequencies were calculated from the number of mutations/total number of bases analyzed and from the total number of bases analyzed from mutated sequences only. NP=nonproductive rearrangement; P=productive rearrangement; ND=Normal Donor *Significant ($p \leq 0.03$) difference in the percentage of mutated NP rearrangements from 4-year-old X-HIgM versus other X-HIgM and ND versus all X-HIgM; †Significant ($p < 0.05$) difference in the overall mutation frequency of NP rearrangements from 4-year-old X-HIgM versus all other X-HIgM and ND versus all X-HIgM; ‡Significant ($p < 0.005$) difference in the mutation frequency of mutated NP rearrangements from ND vs all 4-12 yo X-HIgM and adult X-HIgM vs other X-HIgM.

Table 2. Frequency (%) and distribution of individual nucleotide mutations

	X-HIgM		Normal Donors	
	NP	**P**	**NP**	**P**
RGYW	8.5/51 17%	56.5/250* 23%	147/593* 25%	1227.3/4857* 25%
WRCY	9.5/51 19%	31.5/250 13%	84/593 14%	697.5/4857 14%
WA	8/51† 16%	54/250‡ 22%	137.4/5933‡ 23%	890.5/4857‡ 18%
TW	6/51 12%	22/250 9%	66.7/593 11%	488.5/4857 10%
FR R:S	21/12 1.8	114/50 2.3	304/92§ 3.3	1961/1155 1.7
CDR R:S	14/4 3.5	68/20 3.4	166/29‖ 5.7	1347/394 3.4
Transitions	33/51¶ 65%	140/250 56%	302/593 51%	2626/4857 54%
Transversions	18/51 35%	110/250 44%	291/593 49%	2231/4857 46%

The X-HIgM data include 104 productive (P) and 24 nonproductive (NP) CD19$^+$IgD$^+$ and CD19$^+$CD27$^+$ mutated rearrangements. Normal donor populations include 483 P and 69 NP CD19$^+$IgM$^+$, IgD$^+$CD27$^+$ and IgD$^-$CD27$^+$ mutated rearrangements. R:S= replacement to silent mutation ratio. All positions of the motif were included for mutation analysis. A<u>G</u>CT motifs with a mutated G but not C were considered RGYW motif mutations. AG<u>C</u>T motifs with a mutated C but not G were considered WRCY motif mutations. Three ambiguous A or T mutations in AGCT motifs with neither G or C mutations were scored as a mutation in RGYW and WRCY motifs. *Significant ($p<.05$) difference between RGYW and WRCY mutations; †significant ($p=.03$) difference between X-HIgM and normal donors nonproductive rearrangements; ‡significantly ($p=.05$) greater mutation frequency compared to the expected random frequency; §significant ($p<0.05$) difference between the NP and P FR R:S of ND; ‖significant ($p=0.02$) difference between the NP FR R:S and NP CDR R:S of ND; ¶significant ($p=.03$) difference between the percentage of transitions vs transversions.

Table 3. Target nucleotide mutation frequency (%) in hypermutable motifs

Mutated Base	Motif	X-HIgM		Normal Donor	
		NP	P	NP	P
G	inside RGYW	2/1770 **0.11%**	25.5/7730.5* **0.29%**	92/832* **11.06%**	757/5618* **13.47%**
	all other G	11/12258 **0.09%**	62.5/51099.5 **0.12%**	102/5984 **1.70%**	950/38080 **2.49%**
C	inside WRCY	7.5/1504*† **0.50%**	12.5/6392.5* **0.20%**	60/736* **8.15%**	445/4561* **9.76%**
	all other C	7.5/11034 **0.07%**	39.5/46456 **0.09%**	76/5606 **1.36%**	782/35736 **2.19%**
A	inside WA	8/5370.75 **0.10%**	54/22097* **0.24%**	133.25/2463.25*† **5.41%**	913.8/15111* **6.05%**
	all other A	5/7263.25 **0.07%**	22/30527 **0.07%**	55.75/3412.75 **1.63%**	436.2/22300 **1.96%**
T	inside TW	6/2480* **0.24%**	22/13866* **0.16%**	59.5/1531.25* **3.89%**	498.5/10008.5* **4.98%**
	all other T	5/7788 **0.06%**	12/29179 **0.04%**	44.5/3586.75 **1.24%**	175.5/22503.5 **0.78%**
A	inside RGYW	3.75/1483.25* **0.25%**	15/6545.5* **0.23%**	34/719.25* **4.73%**	291/4701.3* **6.19%**
	all other A	8.25/11150.75 **0.07%**	61/46078.5 **0.13%**	155/5156.75 **3.01%**	1059/32709.7 **3.24%**
A	inside WRCY	1.25/1678.75 **0.07%**	9/7172.5 **0.13%**	27/804.75 **3.36%**	197/4967.8 **3.97%**
	all other A	10.75/10955.25 **0.10%**	67/45451.5 **0.15%**	162/5071.25 **3.19%**	1153/32443 **3.56%**

The data include all nonproductive and productive mutations in X-HIgM patients and normal donors. The targeted nucleotide refers to the base in the underlined position (or two positions, eg. A in the first and second position in the WA motif where W = A or T). The mutation frequency of the targeted nucleotide inside a motif = the number of target nucleotide mutations / the total number of the nucleotide in the motif. The nucleotide mutation frequency outside the motif targeted position(s) = (the total number of specific base mutations in the sequence - the number of targeted nucleotide mutations in the motif) / (the total number of the base in the sequence - the number of target nucleotides inside the motif). The extent of targeting was determined by calculating the fold difference between the mutation frequency inside vs outside a motif. ∗Significant (p<0.02) difference in the nucleotide mutation frequency inside a motif vs outside a motif; †significant (p<0.02) difference in the nucleotide mutation frequency in nonproductive vs productive rearrangements.

Table 4. Gene expression levels in normal donor B cell populations

Cells	Upregulated	Fold change	Downregulated	Fold change	No change
Tonsil					
IgD$^+$CD38^{++}	*AICDA	4.9	CR2	1.5	TNFRSF7 (CD27)
	*APEX2	1.5	*CD40	1.5	POL ζ (XPV)
	*CD38	4.4	MBD4	1.7	POL ζ (Rev3)
	CD80 (B7-1)	1.5	PMSL2 & 3 (PMS)	1.6-2.7	RPA
	*EXO1	4.6	POL μ	2.2	XPA
	*FEN1	2.1-2.2	SMUG1	1.8	XPC
	*MLH1	1.6			XPD
	*MME (CD10)	9.6-13			KU70 (XRCC6)
	*MSH2	1.6			KU80 (XRCC5)
	MSH3	1.5			LIG4
	MSH5	1.4-1.8			
	*MSH6	1.5-1.9			
	*NEIL1	2.9			
	*OGG1	2.0			
	*PCNA	2.5			
	*POL θ	2.1-2.6			
	*RAD50	1.4			
	*RAD51	4.2-6.5			
	*UNG	1.8			
	UNG2	1.7			
	*XRCC4	1.0-5.2			
Tonsil					
IgD$^-$CD38^{++}	*AICDA	15.4	*CR2	1.7	POL ζ (Rev3)
	*CD38	8.8	*ERCC1	1.5	XPA
	*CD80 (B7-1)	1.8	*MSH4	1.5	XPC
	*EXO1	6.4	POL ι	1.5	XPD
	*FEN1	2.8-3.5	POL μ	3.3	KU70 (XRCC6)
	*MLH1	1.8	POL η	2.7	KU80 (XRCC5)
	*MME (CD10)	35-44			LIG4
	*MSH2	2.1			
	*PCNA	3.7			
	*POL θ	1.6-3.7			
	*RAD51	6.2-11.9			
	*RPA	1.8			
	*TNFRSF7 (CD27)	3.1			
	*UNG	3.0			

Tonsil populations were determined from 3 individual normal donors. *Significant (p <0.05) differences between tonsillar naïve B cells (CD38$^-$IgD$^+$) versus activated naïve/pre-GC founder cells (CD38^{++}IgD$^+$) or tonsillar GC cells (CD38^{++}IgD$^-$).

Figure 1

A

B

Figure 2

Percentage of CD19+ B cells expressing IgD and CD27
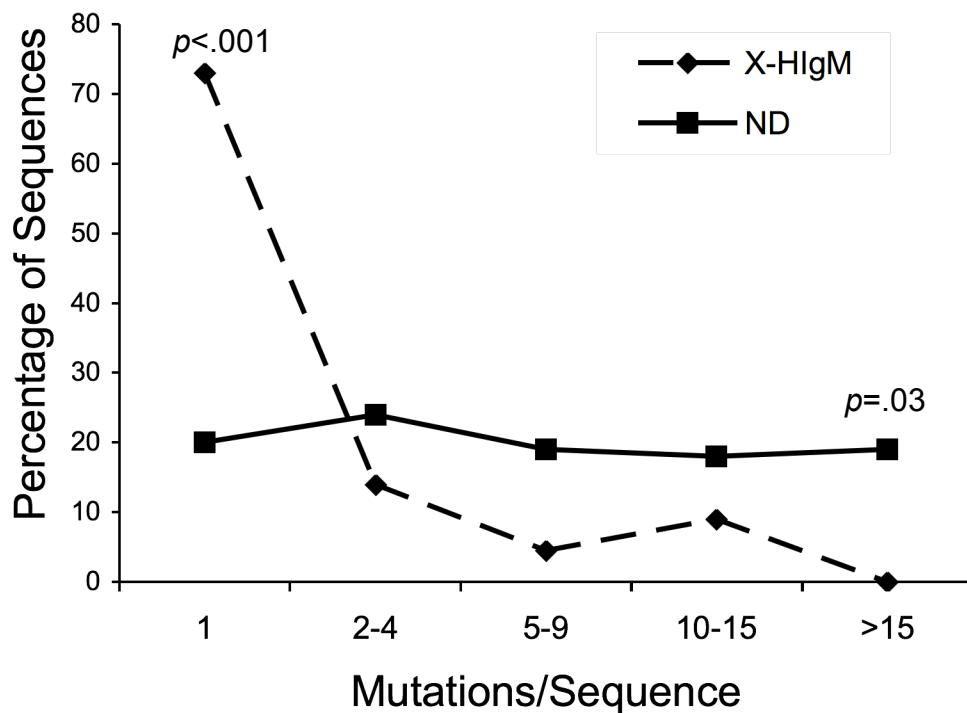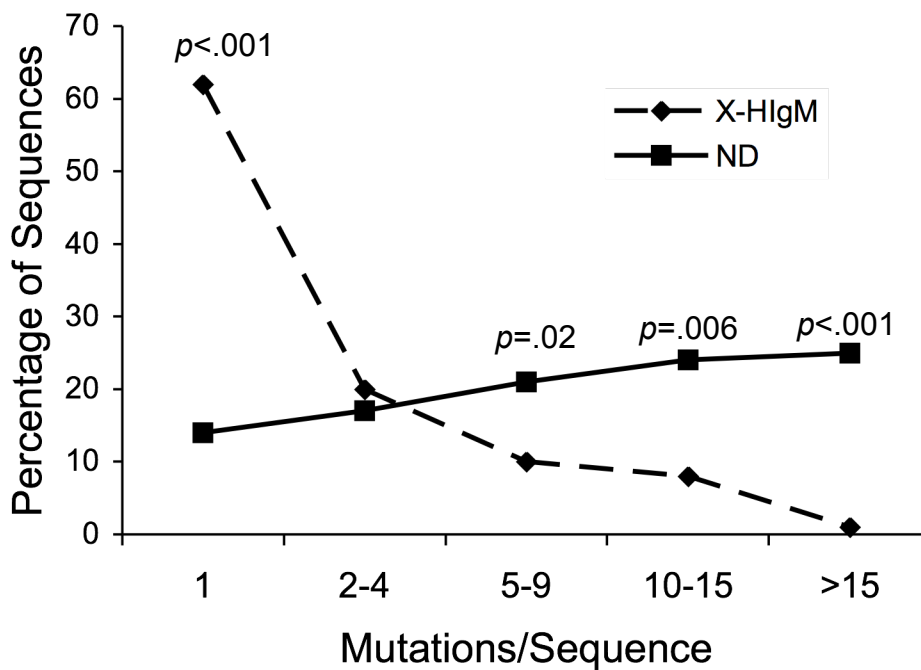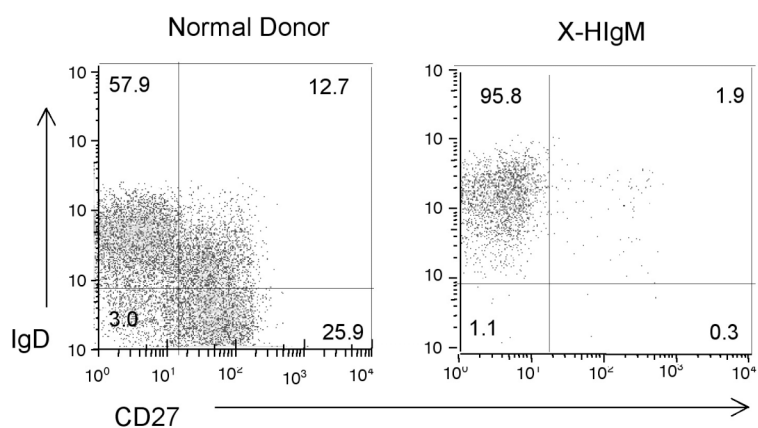
	IgD+CD27-	IgD+CD27+	IgD-CD27+
ND Adult	57.0*	12.7*	25.9*
X-HIgM #1	97.1	1.3	0.6
X-HIgM #2	97.3	1.1	0.6
X-HIgM #3	96.6	2.5	0.7
X-HIgM #4	95.8	1.9	0.3

*Significant ($p<0.001$) difference between ND and X-HIgM

PAGE   6

Figure 3

A. Nonproductive Rearrangements

X-HIgM

Number of mutations from	To	A	C	G	T		total
	A	-	1	6	5		12
	C	1	-	1	13		15
	G	8	2	-	3		13
	T	3	6	2	-		11
							51

Normal Donors

	To	A	C	G	T		total
	A	-	49	92	48		189‡
	C	24	-	39	60		123
	G	99	65	-	25		189†
	T	21	51	20	-		92
							593

Percent mutations from	To	A	C	G	T		total
	A	-	2	12	10		24
	C	2	-	2	25		29
	G	14	3	-	5		22
	T	7	14	5	-		25
							100

	To	A	C	G	T		total
	A	-	9	16	8		33
	C	4	-	6	10		20
	G	15	10	-	4		29
	T	4	10	4	-		18
							100

B. Productive Rearrangements

X-HIgM

Number of mutations from	To	A	C	G	T		total
	A	-	13	40	23		76‡
	C	3	-	10	36*		49
	G	45	29	-	15		89†
	T	8	19	9	-		36
							250

Normal Donors

	To	A	C	G	T		total
	A	-	304	700	309		1313‡
	C	180	-	311	709		1200
	G	840	606	-	231		1677†
	T	155	377	135	-		667
							4857

Percent mutations from	To	A	C	G	T		total
	A	-	5	16	9		30
	C	1	-	4	14		19
	G	16	10	-	5		31
	T	4	9	4	-		17
							97

	To	A	C	G	T		total
	A	-	6	15	7		28
	C	4	-	6	14		24
	G	15	11	-	4		30
	T	4	9	3	-		16
							98

PAGE   7